10th Workshop on Statistical Machine Translation 2015

Held at EMNLP 2015

Lisbon, Portugal
17 – 18 September 2015

ISBN: 978-1-5108-1188-1
### Table of Contents

**Findings of the 2015 Workshop on Statistical Machine Translation**
Ondřej Bojar, Rajen Chatterjee, Christian Federmann, Barry Haddow, Matthias Huck, Chris Hokamp, Philipp Koehn, Varvara Logacheva, Christof Monz, Matteo Negri, Matt Post, Carolina Scarton, Lucia Specia and Marco Turchi .......................................................... 1

**Statistical Machine Translation with Automatic Identification of Translationese**
Naama Twitto, Noam Ordan and Shuly Wintner ........................................... 47

**Data Selection With Fewer Words**
Amittai Axelrod, Philip Resnik, Xiaodong He and Mari Ostendorf .................... 58

**DFKI’s experimental hybrid MT system for WMT 2015**
Eleftherios Avramidis, Maja Popović and Aljoscha Burchardt .......................... 66

**ParFDA for Fast Deployment of Accurate Statistical Machine Translation Systems, Benchmarks, and Statistics**
Ergun Bicici, Qun Liu and Andy Way ......................................................... 74

**CUNI in WMT15: Chimera Strikes Again**
Ondřej Bojar and Aleš Tamchyna .............................................................. 79

**CimS - The CIS and IMS Joint Submission to WMT 2015 addressing morphological and syntactic differences in English to German SMT**
Fabienne Cap, Marion Weller, Anita Ramm and Alexander Fraser .................... 84

**The Karlsruhe Institute of Technology Translation Systems for the WMT 2015**
Eunah Cho, Thanh-Le Ha, Jan Niehues, Teresa Herrmann, Mohammed Mediani, Yuqi Zhang and Alex Waibel ......................................................... 92

**New Language Pairs in TectoMT**
Ondřej Dušek, Luís Gomes, Michal Novák, Martin Popel and Rudolf Rosa .......... 98

**Tuning Phrase-Based Segmented Translation for a Morphologically Complex Target Language**
Stig-Arne Grönroos, Sami Virpioja and Mikko Kurimo .................................... 105

**The AFRL-MITLL WMT15 System: There’s More than One Way to Decode It!**
Jeremy Gwinnup, Tim Anderson, Grant Erdmann, Katherine Young, Christina May, Michael Kazi, Elizabeth Salesky and Brian Thompson ............................ 112

**The KIT-LIMSI Translation System for WMT 2015**
Thanh-Le Ha, Quoc-Khanh DO, Eunah Cho, Jan Niehues, Alexandre Allauzen, François Yvon and Alex Waibel ......................................................... 120

**The Edinburgh/JHU Phrase-based Machine Translation Systems for WMT 2015**
Barry Haddow, Matthias Huck, Alexandra Birch, Nikolay Bogoychev and Philipp Koehn ................................................................. 126

**Montreal Neural Machine Translation Systems for WMT’15**
Sébastien Jean, Orhan Firat, Kyunghyun Cho, Roland Memisevic and Yoshua Bengio .............. 134

**GF Wide-coverage English-Finnish MT system for WMT 2015**
Prasanth Kolachina and Aarne Ranta .......................................................... 141
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIMSI@WMT’15 : Translation Task</td>
<td>Benjamin Marie, Alexandre Allauzen, Franck Burlot, Quoc-Khanh Do, Julia Ive, elena knyazeva, Matthieu Labeau, Thomas Lavergne, Kevin Löser, Nicolas Pécheux and François Yvon</td>
<td>145</td>
</tr>
<tr>
<td>UdS-Sant: English–German Hybrid Machine Translation System</td>
<td>Santanu Pal, Sudip Naskar and Josef van Genabith</td>
<td>152</td>
</tr>
<tr>
<td>The RWTH Aachen German-English Machine Translation System for WMT 2015</td>
<td>Jan-Thorsten Peter, Farzad Toutounchi, Joern Wuebker and Hermann Ney</td>
<td>158</td>
</tr>
<tr>
<td>Exact Decoding with Multi Bottom-Up Tree Transducers</td>
<td>Daniel Quernheim</td>
<td>164</td>
</tr>
<tr>
<td>Sheffield Systems for the Finnish-English WMT Translation Task</td>
<td>David Steele, Karin Sim Smith and Lucia Specia</td>
<td>172</td>
</tr>
<tr>
<td>Morphological Segmentation and OPUS for Finnish-English Machine Translation</td>
<td>Jörg Tiedemann, Filip Ginter and Jenna Kanerva</td>
<td>177</td>
</tr>
<tr>
<td>The University of Illinois submission to the WMT 2015 Shared Translation Task</td>
<td>Lane Schwartz, Bill Bryce, Chase Geigle, Sean Massung, Yisi Liu, Haoruo Peng, Vignesh Raja, Subhro Roy and Shyam Upadhyay</td>
<td>192</td>
</tr>
<tr>
<td>Edinburgh’s Syntax-Based Systems at WMT 2015</td>
<td>Philip Williams, Rico Sennrich, Maria Nadejde, Matthias Huck and Philipp Koehn</td>
<td>199</td>
</tr>
<tr>
<td>The FBK Participation in the WMT15 Automatic Post-editing Shared Task</td>
<td>Rajen Chatterjee, Marco Turchi and Matteo Negri</td>
<td>210</td>
</tr>
<tr>
<td>Why Predicting Post-Edit is so Hard? Failure Analysis of LIMSI Submission to the APE Shared Task</td>
<td>Guillaume Wisniewski, Nicolas Pécheux and François Yvon</td>
<td>222</td>
</tr>
<tr>
<td>Hierarchical Machine Translation With Discontinuous Phrases</td>
<td>Miriam Kaeshammer</td>
<td>228</td>
</tr>
<tr>
<td>Discontinuous Statistical Machine Translation with Target-Side Dependency Syntax</td>
<td>Nina Seemann and Andreas Maletti</td>
<td>239</td>
</tr>
<tr>
<td>ListNet-based MT Rescoring</td>
<td>Jan Niehues, Quoc-Khanh DO, Alexandre Allauzen and Alex Waibel</td>
<td>248</td>
</tr>
<tr>
<td>Results of the WMT15 Metrics Shared Task</td>
<td>Miloš Stanojević, Amir Kamran, Philipp Koehn and Ondřej Bojar</td>
<td>256</td>
</tr>
<tr>
<td>Results of the WMT15 Tuning Shared Task</td>
<td>Miloš Stanojević, Amir Kamran and Ondřej Bojar</td>
<td>274</td>
</tr>
<tr>
<td>Extended Translation Models in Phrase-based Decoding</td>
<td>Andreas Guta, Joern Wuebker, Miguel Graca, Yunsu Kim and Hermann Ney</td>
<td>282</td>
</tr>
</tbody>
</table>
Investigations on Phrase-based Decoding with Recurrent Neural Network Language and Translation Models
  Tamer Alkhouli, Felix Rietig and Hermann Ney .......................................................... 294

Referential Translation Machines for Predicting Translation Quality and Related Statistics
  Ergun Bicici, Qun Liu and Andy Way ................................................................. 304

UAlacant word-level machine translation quality estimation system at WMT 2015
  Miquel Esplà-Gomis, Felipe Sánchez-Martínez and Mikel Forcada ......................... 309

QUality Estimation from ScraTCH (QUETCH): Deep Learning for Word-level Translation Quality Estimation
  Julia Kreutzer, Shigehiko Schamoni and Stefan Riezler ....................................... 316

LORIA System for the WMT15 Quality Estimation Shared Task
  David Langlois ...................................................................................................... 323

Data enhancement and selection strategies for the word-level Quality Estimation
  Varvara Logacheva, Chris Hokamp and Lucia Specia .......................................... 330

USHEF and USAAR-USHEF participation in the WMT15 QE shared task
  Carolina Scarton, Liling Tan and Lucia Specia ................................................... 336

SHEF-NN: Translation Quality Estimation with Neural Networks
  Kashif Shah, Varvara Logacheva, Gustavo Paetzold, Frédéric Blain, Daniel Beck, Fethi Bougares and Lucia Specia ............................................................. 342

Strategy-Based Technology for Estimating MT Quality
  Liugang Shang, Dongfeng Cai and Duo Ji .................................................................. 348

UGENT-LT3 SCATE System for Machine Translation Quality Estimation
  Arda Tezcan, Veronique Hoste, Bart Desmet and Lieve Macken ............................. 353

Multi-level Evaluation for Machine Translation
  Boxing Chen, Hongyu Guo and Roland Kuhn ........................................................ 361

VERTa: a Linguistically-motivated Metric at the WMT15 Metrics Task
  Elisabet Comelles and Jordi Atserias ...................................................................... 366

UPF-Cobalt Submission to WMT15 Metrics Task
  Marina Fomicheva, Núria Bel, Iria da Cunha and Anton Malinovskiy ...................... 373

Machine Translation Evaluation using Recurrent Neural Networks
  Rohit Gupta, Constantin Orasan and Josef van Genabith ....................................... 380

Alignment-based sense selection in METEOR and the RATATOUILLE recipe
  Benjamin Marie and Marianna Apidianaki ............................................................ 385

chrF: character n-gram F-score for automatic MT evaluation
  Maja Popović ........................................................................................................ 392

BEER 1.1: ILLC UvA submission to metrics and tuning task
  Miloš Stanojević and Khalil Sima’an ..................................................................... 396

Predicting Machine Translation Adequacy with Document Embeddings
  Mihaela Vela and Liling Tan ................................................................................. 402